基于深度学习语义分割和迁移学习策略的麦田倒伏面积识别方法

张　淦[1]，严海峰[1]，胡根生[1*]，张东彦[1,2]，程　涛[1,2]，潘正高[1,3]，许海峰[1,3]，沈书豪[1,3]，朱科宇[1]

（1. 安徽大学农业生态大数据分析与应用技术国家地方联合工程研究中心，安徽合肥 230039，中国；2. 西北农林科技大学 机械与电子工程学院，陕西杨凌 712100，中国；3. 宿州学院 信息工程学院，安徽宿州 234000，中国）

摘　要：［**目的/意义**］利用低空无人机技术并结合深度学习语义分割模型精准提取作物倒伏区域是一种高效的倒伏灾害监测手段。然而，在实际应用中，受田间各种客观条件（不同无人机飞行高度低于120 m、多个研究区、关键生育期不同天气状况等）限制，无人机获取的图像数量仍偏少，难以满足高精度深度学习模型训练的要求。本研究旨在探索一种在作物生育期和研究区有限的情况下精准提取倒伏面积的方法。［**方法**］以健康/倒伏小麦为研究对象，在其灌浆期和成熟期开展麦田图像采集工作。设置2个飞行高度（40和80 m），采集并拼接获取2019、2020、2021和2023年份3个研究区的数字正射影像图（Digital Ortho-photo Map，DOM）；在Swin-Transformer深度学习语义分割框架基础上，分别使用40 m训练集单独训练、40和80 m训练集混合训练、40 m训练集预训练80 m训练集迁移学习等3种训练方法，获得对照模型、混合训练模型和迁移学习模型；采用对比实验比较上述3种模型分割80 m高度预测集图像的精度并评估模型性能。［**结果和讨论**］迁移学习模型倒伏面积提取精度最高，交并比、正确率、精确率、召回率和F_1-Score共5个指标平均数分别为85.37%、94.98%、91.30%、92.52%和91.84%，高于对照组模型1.08%～3.19%，平均加权帧率达到738.35 fps/m^2，高于40 m图像183.12 fps/m^2。［**结论**］利用低飞行高度（40 m）预训练语义分割模型，在较高飞行高度（80 m）空图像做迁移学习的方法提取倒伏小麦面积是可行的，这为解决空域飞行高度限制下，较少80 m及以上图像数据集无法满足语义分割模型训练的要求的问题，提供了一种有效的方法。

关键词：倒伏识别；农业遥感；无人机影像；迁移学习；语义分割；Swin-Transformer

中图分类号：S46;TP183　　**文献标志码：**A　　**文章编号：**SA202309013

引用格式：张淦，严海峰，胡根生，张东彦，程涛，潘正高，许海峰，沈书豪，朱科宇. 基于深度学习语义分割和迁移学习策略的麦田倒伏面积识别方法[J]. 智慧农业(中英文)，2023, 5(3): 75-85. DOI：10.12133/j.smartag.SA202309013

ZHANG Gan, YAN Haifeng, HU Gensheng, ZHANG Dongyan, CHENG Tao, PAN Zhenggao, XU Haifeng, SHEN Shuhao, ZHU Keyu. Identification method of wheat field lodging area based on deep learning semantic segmentation and transfer learning[J]. Smart Agriculture, 2023, 5(3): 75-85. DOI：10.12133/j.smartag.SA202309013 (in Chinese with English abstract)

1　引　言

2022年国家统计数据显示，中国小麦产量占据粮食总产量的20.06%，仅次于稻谷和玉米，在保证国家粮食安全方面具有至关重要的地位[1]。但是，受到极端气象灾害（如台风、暴雨、干旱、洪涝等）的影响，小麦容易产生倒伏，引发各种次生病

收稿日期：2023-09-11

基金项目：安徽省教育厅高校科研项目（自然科学类）（2023AH052246）；宿州学院博士科研启动基金（2021BSK043）；国家自然科学基金（42271364）

作者简介：张　淦，研究方向为农业遥感。E-mail：zhanggan@ahu.edu.cn

*通信作者：胡根生，博士，教授，研究方向为机器学习及图像视觉。E-mail：hugs2906@sina.com

害，导致小麦的产量和品质降低[2-4]。倒伏严重程度调查是倒伏灾害预警和防治工作的重要环节，随着农业遥感技术的发展，倒伏监测的精度和效率得到大幅提升[5-7]。传统的作物倒伏遥感监测主要包括地面、航空和卫星三个尺度。Chauhan等[8]利用车载传感器获取了稻田图像，并通过分析可见光和近红外波段的光谱信息，成功地预测了稻田倒伏程度。李宗南等[9]在其研究中选择了Worldview-2卫星图像中的红色波段、近红外波段1和近红外波段2作为首选波段，然后采用最大似然分类方法提取了玉米倒伏面积，平均误差率为4.7%。Yang等[10]使用合成孔径雷达卫星图像对麦田倒伏情况进行监测，提出了旋光指数法，几乎所有的倒伏区都成功地与正常区区分开来。相比于卫星影像，无人机影像分辨率较高，受到地形和大气干扰较小，对于倒伏特征还原性更佳，常用于高精度倒伏面积提取。然而，无人机倒伏监测常受作业面积大小和作物关键生育期天气状况（大风、降雨等）的限制，获得的有效数据仍然偏少，无法满足深度学习模型的样本量需求[11-13]，导致提出的方法普适性仍欠佳。迁移学习是一种解决样本数量不足的方法[14]，可用于提高无人机倒伏监测模型的普适性。

当前，深度学习技术在农业领域取得了显著成果，在农业领域的应用涵盖了农作物病害评估、农作物倒伏监测、品种分类、杂草识别以及农作物计数等。而迁移学习技术是一种常用的深度学习训练方法，在农业目标检测和图像分割模型训练中得到了广泛的应用。Zhao等[15]采用了UNet网络结构，运用迁移学习策略成功提取了晚熟期水稻倒伏面积，dice系数达到了0.94。Yang等[16]使用RGB图像，运用迁移学习策略训练FCN-AlexNet网络来提取水稻倒伏面积，其准确度达到了0.94。为了进一步提高倒伏面积提取的准确性，一些研究者进行了深度学习算法的创新。Zhao等[17]基于PSPNet网络，采用了轻量级神经网络MobileNetV2来替代传统的ResNet作为特征提取骨干网络，并使用深度可分离卷积代替标准卷积，并使用迁移学习策略冻结部分网络层进行训练，改进后的PSPNet提取小麦倒伏面积性能更优。Zhang等[18]利用迁移学习训练DeepLabv3+模型，分割不同生育期的小麦倒伏区域，使用迁移学习方法的效果优于传统的UNet方法。Yu等[19]引入了卷积算子和Dense block模块对UNet网络进行改进，使用迁移学习策略训练模型，并利用不同飞行高度采集的数据进行倒伏面积提取，结果表明，20 m飞行高度的分割效果最佳。当前流行的图像分割算法模型多基于深度学习语义分割技术，模型的性能与模型的参数量呈现正相关关系，如UNet[20]的参数量达到了2800万个，DeepLabV3[21]的参数量达到了4000万个，而目前功能最强大的语义分割模型Swin-Transfomer[22]模型参数量达到了30亿个，训练样本量不足将会导致模型出现过拟合。

综上发现，已有的深度学习算法适用于倒伏面积提取模型开发，但仍然存在如下瓶颈问题制约其推广应用价值：（1）模型泛化性能未得到充分验证。上述研究获取的无人机图像年份以及研究区较少，样本数量不足，采用的模型多为低容量的轻量级深度学习网络模型，模型对于其他农业场景的适应性未得到充分验证；（2）倒伏监测效率低下。研究多采用20和40 m高度采集的无人机图像，农田测绘耗时较长，消耗大量人工，不利于技术的推广。

针对上述问题，本研究使用大型语义分割模型Swin-Transformer框架，针对2019、2020、2021和2023年采集的3个研究区40和80 m高度的麦田低空遥感图像开展倒伏面积提取算法研究，预测集和训练集采用不同年份的图像。为了克服80 m高度采集的图像样本数量不能满足Swin-Transformer训练要求的问题，提出了利用40 m图像预训练语义分割模型，再用少量80 m图像做迁移学习的策略，并通过对比实验的方法进行验证，以期为作物倒伏监测提供新方法和新思路。

2 材料和方法

2.1 研究区介绍

研究区位于安徽省中部合肥市，分别是舒城县千人桥镇山林种植专业合作社（117.1°E，31.5°N，以下简称“舒城”）、庐江县郭河镇国家现代农业示范区（117.2°E，31.5°N，以下简称“郭河”）和

庐江县白湖镇小麦综合实验站（117.5°E，31.2°N，以下简称“白湖”）。研究区地形平坦，属于温带大陆性气候，年平均气温约13℃，年均降水量约为1480 mm，年平均风速约10.7 km/h，种植制度主要为“冬小麦-大豆/玉米”一年两熟。其中郭河采用条播，地块长度约150 m，宽度约3 m，播种带间距约0.3 m，而白湖和舒城采用撒播。根据合肥市气象局数据结合田间调查结果，分别在2019年5月14日至5月18日，2020年5月1日、5月5日、5月8日和5月14日，2021年5月11日至5月15日、5月23日，2023年5月2日到5月6日，研究区出现较强降水，引起大面积的小麦倒伏。研究区域地理位置及现场倒伏情况见图1和图2。



(a-1)舒城研究区现场



(b-1)郭河研究区现场



(c-1)白湖研究区现场



(a-2)舒城研究区测绘图



(b-2)郭河研究区测绘图



(c-2)白湖研究区测绘图

图1 舒城、郭河和白湖研究区现场图和测绘图

Fig. 1 Site maps and surveying maps of the study areas in Shucheng, Guohe and Baihu



(a)灌浆期小麦倒伏



(b)成熟期小麦倒伏

图2 灌浆期和成熟期小麦倒伏

Fig. 2 Wheat lodging during the filling and mature period

2.2 图像采集

于2019、2020、2021和2023年在相同的麦田拍摄无人机影像，图像采集时间为每年4月底至6月初，这里只列出小麦灌浆期和成熟期两个生育期。图像采集时刻为北京时间10：00—15：00。使用大疆精灵4RTK和Mavic 3E四轴专业级无人机采集麦田图像，均搭载2000万像素高清摄像机，具备测绘功能。无人机飞行高度被设定为40和80 m，相对应的空间分辨率约为2.0和4.1 cm。相机拍照模式选择等距间隔拍照，航向重叠率为75%，旁向重叠率为70%，云台俯仰角度为−90°，飞行速度为6 m/s。

图像采集完成后，使用“大疆智图”软件（版

本号3.6.0)对航拍图片进行拼接，生成各研究区数字正射影像图。其中，40 m飞行高度下图像分辨率约为2000万～6000万像素，80 m高度图像分辨率约为500万～1500万像素，并请农学专家依据田间调查结果标注倒伏区域。舒城、郭河和白湖研究区麦田图像信息见表1。

表1 舒城、郭河和白湖研究区麦田图像信息统计

Table1 Statistics information of wheat fields of the study areas in Shucheng, Guohe and Baihu

研究区	年份	作业面积/m^2	图像采集天数/d	拼接图像数量/张		倒伏面积占比/%
				40 m飞行高度	80 m飞行高度	
舒城	2023	33,041.5	5	5	5	13.84
郭河	2020	13,521.2	3	3	——	19.81
	2021	13,446.4	4	4	3	28.11
	2023	13,907.4	5	5	5	64.99
白湖	2019	28,451.2	4	4	1	6.52
	2020	28,553.4	2	3	——	14.81
	2021	28,556.3	3	3	3	11.61
	2023	28,573.7	5	5	5	17.93

2.3 研究方法

本研究利用迁移学习策略解决样本不足问题。无人机在40 m飞行高度时采集的图像(以下简称“40 m图像”)空间分辨率较高，每次航拍采集到的图像数量多于80 m飞行高度采集的图像(以下简称“80 m图像”)，且40 m图像可以近似看作为80 m图像的上采样，图像中的倒伏特征较为接近。因此，利用40 m图像预训练语义分割模型，再用少量80 m图像做迁移学习的策略，是一种解决80 m图像数量无法满足语义分割模型训练要求问题的可行性方案。使用对比实验的方法验证上述策略的有效性，技术路线如图3所示。

2.3.1 实验准备

本研究在2个飞行高度(40和80 m)下采集了多年份(2019、2020、2021和2023年)、不同区域(3个研究区)、小麦关键生育期(灌浆期和成熟期)的麦田倒伏数据。将拼接后的图像按照高度和年份划分为预训练数据集(2019、2020、2021年40 m图像)、迁移学习数据集(2019、2020、2021年80 m图像)和预测数据集(2023年40和80 m图像)。其

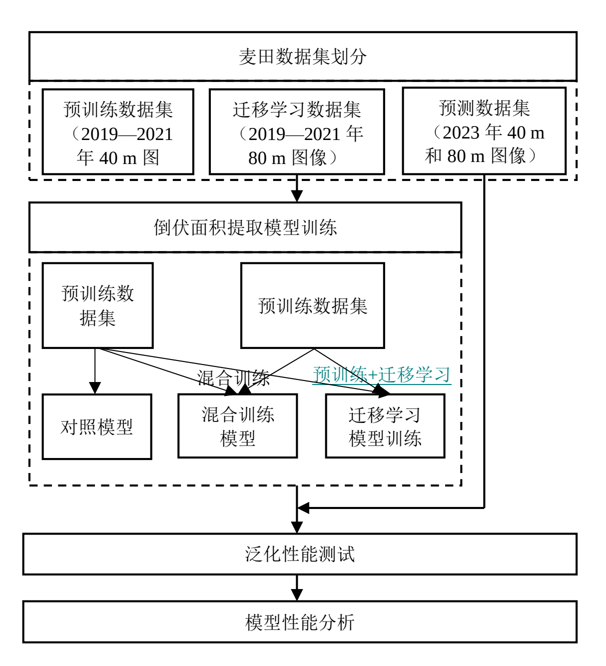


图3 小麦倒伏区域提取迁移学习模型研究技术路线

Fig. 3 Technical route for wheat lodging areas detection based on transfer learning

中，预训练数据集拼接后图像共17幅，迁移学习数据集拼接后图像共7幅，预测数据集拼接后图像共30幅。相比于其他倒伏面积提取算法研究[15-19]，本研究的数据较为充足。

根据前期研究成果[23]，针对多年份单一无人机飞行高度采集的图像，Swin-Transformer模型获得的分割精度比U-Net、SegNet、DeepLab等基于卷积神经网络的语义分割模型高5%～10%。因此，选择Swin-Transformer[22]模型框架用于训练倒伏面积提取模型。

2.3.2 模型训练方法

为验证本研究提出的迁移学习策略的有效性，基于Swin-Transformer框架，使用2个飞行高度下4个年份3个研究区的无人机图像，通过训练获得对照模型、混合训练模型和迁移学习训练模型。使用预训练数据集训练网络模型，得到对照模型；在对照模型的基础上，使用迁移学习数据集进行迁移学习，获得迁移学习模型；将预训练数据集和迁移学习数据集混合，共同训练网络模型，获得混合训练数据集。

2.3.3 模型训练参数设置

硬件运行环境为AMD Ryzen 7 5800X 8-core processor CPU服务器，内存为64 G，搭载3080ti显

卡，显存12 G，Windows10系统，采用Pytorch编写深度学习算法。

为验证利用40 m飞行高度图像预训练语义分割模型，再用少量80 m高度图像做迁移学习的策略对于倒伏面积提取任务的有效性，设置了3组对比实验：基于Swin-Transformer框架，训练对照模型、混合训练模型和迁移学习模型。

2.3.4　评价指标

倒伏面积提取任务属于语义分割中的二分类问题，即将图像像素分割为倒伏和非倒伏两类。对于模型性能的评估，本研究采用经典的语义分割模型精度评估指标体系，共5个指标，交并比（Intersection of Union，IOU）、正确率（Accuracy）、精确率（Precision）、召回率（Recall）和F_1-Score。具体计算如公式（1）～（5）。

$$IOU = TP/(FP + TP + FN) \tag{1}$$

$$Accuracy = (TP + TN)/(TP + FP + TN + FN) \tag{2}$$

$$Precision = TP/(TP + FP) \tag{3}$$

$$Recall = TP/(TP + FN) \tag{4}$$

$$F_1\text{-}Score = (1+\beta^2)/\left(\left(\frac{1}{\beta^2 \times Precision}\right)+\left(\frac{1}{Recall}\right)\right) \tag{5}$$

其中，结合研究内容，TP、FP、TN、FN的在本研究的定义见混淆矩阵（表2）。

表2　混淆矩阵

Table 2　Confusion matrix

混淆矩阵		图像标注结果	
		倒伏	非倒伏
图像分割结果	倒伏	TP	FP
	非倒伏	FN	TN

具体而言，对于交并比、精确度、召回率和F_1-Score这4个指标，倒伏和非倒伏区域均可以作为正类。通过对多个年份不同研究区倒伏状况进行统计，发现不同研究区和生育期麦田倒伏和非倒伏面积的比例差异较大，正确率指标不能全面地反映算法对于倒伏区域识别的精度。因此，分别将倒伏和非倒伏区域作为正类，计算交并比、精确度、召回率和F_1-Score，并求取平均值，与正确率指标共同作为评估模型性能的指标。

本研究采用的模型为Swin-Transformer，图像分割时间消耗与图像的像素数量呈正相关，图像像素又与无人机的飞行高度和研究区的面积相关，其中，无人机飞行高度是研究的变量，而研究区的面积是干扰因素。因此，为了更加客观地比较40和80 m飞行高度下的算法效率，应当去除研究区面积的干扰。本研究使用遥感图像算法效率比较指标——加权帧率wf，即将图像处理的帧率乘以研究区面积（公式（6））进行比较确定。

$$wf = fr \times area \tag{6}$$

其中，fr是图像处理帧率，f/s；$area$是研究区面积，m^2。

2.3.5　泛化性能测试

泛化性能是机器学习中的一个重要指标，它指的是模型在新数据上的表现能力。泛化性能的提升是模式识别领域的重要研究方向，也是评估模型性能的重要标准之一。针对倒伏识别问题，模型对于不同研究区农业非结构化环境的适应能力，决定了模型提取真实农业场景下倒伏区域的效果。因此，本研究以小麦倒伏监测模型的泛化性能作为评估模型性能的指标之一。通过对比对照模型，混合训练模型和迁移学习模型的泛化性能相关精度指标，验证利用低尺度空图像预训练语义分割模型，再用少量较高尺度图像做迁移学习的策略对于倒伏面积提取任务的有效性。

预测集的数据被按照无人机的飞行高度划分为40 m预测集和80 m预测集两部分。比较各模型对2个预测集图像倒伏区域提取的精度，有助于测试迁移学习策略是否使得各模型充分学习80 m图像倒伏区域特征。更深一步，每个子数据集依据所属研究区被分割3部分。比较3个模型对于同一研究区获取的图像倒伏区域提取的精度，有助于克服采集地点不同造成的倒伏特征系统性差异；而比较同一模型对于3个研究区获取的图像倒伏区域提取的精度，更能体现模型的泛化性能。

3　结果与分析

无人机在40和80 m飞行高度下采集的倒伏图像空间分辨率存在差异。本研究重点揭示这种差异对于倒伏建模算法精度产生的影响以及迁移学习策略对于提升模型精度的效果。其中，表3是对照模

型、混合训练模型和迁移学习模型对预测集数据中倒伏面积的提取结果，表4是各模型的平均加权帧率的计算结果，图4是倒伏面积提取结果。

表3 舒城、郭河和白湖研究区倒伏面积提取模型精度和算法速度

Table 3 Accuracy and speed for lodging area detection of the study areas in Shucheng, Guohe and Baihu

模型	飞行高度/m	研究区	交并比/%	正确率/%	精确率/%	召回率/%	F_1-Score/%	算法速度/(f·s^{-1})
对照模型	40	舒城	86.78	97.21	90.80	94.65	92.61	0.0055
		郭河	81.03	90.15	89.15	89.81	89.45	0.0122
		白湖	90.47	97.19	94.40	95.38	94.88	0.0064
	80	舒城	82.40	96.53	90.20	89.24	89.72	0.0228
		郭河	75.40	86.68	85.02	88.69	85.88	0.0504
		白湖	88.74	96.70	95.43	92.47	93.87	0.0265
混合训练模型	40	舒城	84.34	96.45	87.75	95.32	91.06	0.0053
		郭河	80.11	89.84	89.52	88.33	88.86	0.0118
		白湖	88.64	96.47	91.76	96.28	93.82	0.0061
	80	舒城	84.35	96.79	89.29	93.00	91.04	0.0231
		郭河	81.84	90.75	88.95	91.53	89.94	0.0525
		白湖	89.61	96.94	95.34	93.50	94.39	0.0267
迁移学习模型	40	舒城	84.29	96.40	87.42	95.80	91.03	0.0057
		郭河	81.16	90.49	90.56	88.72	89.51	0.0124
		白湖	88.36	96.36	91.48	96.30	93.66	0.0066
	80	舒城	83.72	96.68	89.33	92.03	90.62	0.0227
		郭河	82.71	91.32	89.66	91.56	90.46	0.0511
		白湖	89.68	96.93	94.90	93.97	94.43	0.0264

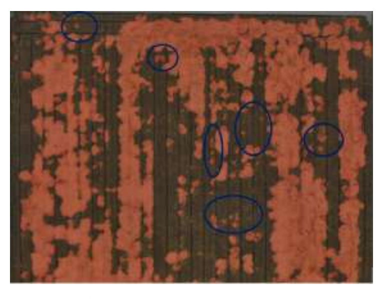
表4 舒城、郭河和白湖研究区各模型平均加权帧率

Table 4 Mean weighted frame rate of each model of the study areas in Shucheng, Guohe and Baihu

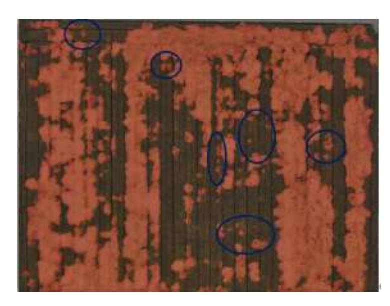
模型	飞行高度/m	平均加权帧率/(fps·m^{-2})
对照模型	40	178.09
	80	737.16
混合训练模型	40	171.18
	80	752.10
迁移学习模型	40	183.12
	80	738.35

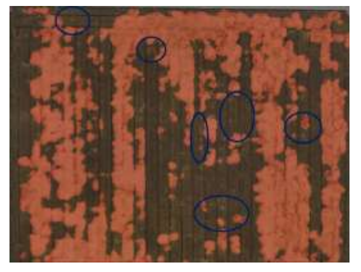
3.1 飞行高度对小麦倒伏面积提取平均精度影响分析

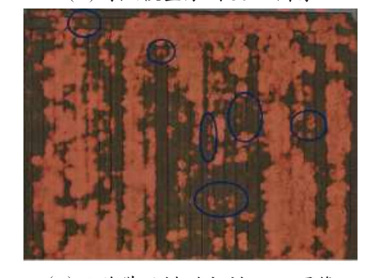
结合表3和图4，当只使用40 m高度图像训练，对照模型分割40 m高度图像的精度优于80 m图像，3个研究区80 m预测集图像倒伏面积提取精度评价指标交并比、正确率、精确率、召回率和F_1-Score的平均数分别为82.18%、93.30%、90.22%、90.13%和89.82%，而40 m预测集上述5个精度指标平均数分别为86.09%、94.85%、91.45%、93.28%和92.31%，40 m预测集图像的5个指标分别高于80 m图像3.91%、1.55%、1.23%、3.15%和2.49%。当使用混合训练和迁移学习策略，加入80 m高度数据训练模型时，模型对于80 m高度图像中倒伏面积的提取结果上升，但是对于40 m图像提取精度下降，使得混合训练模型和迁移学习模型对40和80 m图像中倒伏面积提取精度较为接近。其中，混合训练模型80 m预测集图像倒伏面积提取精度评价指标交并比、正确率、精确率、召回率和F_1-Score的平均数分别为85.27%、94.83%、91.19%、92.68%和91.79%，40 m预测集图像上述5个精度指标分别为84.36%、94.25%、89.68%、93.31%和91.25%，迁移学习模型80 m预测集图像倒伏面积提取精度评价指标交并比、正确率、精确率、召回率和F_1-Score的平均数为85.37%、94.98%、91.30%、92.52%和91.84%，40 m预测集图像的上述5个精度指标分别为84.60%、94.42%、89.82%、93.61%和91.40%。


(a)对照模型分割40 m图像


(b)对照模型分割80 m图像


(c)混合训练模型分割40 m图像


(d)混合训练模型分割80 m图像


(e)迁移学习模型分割40 m图像

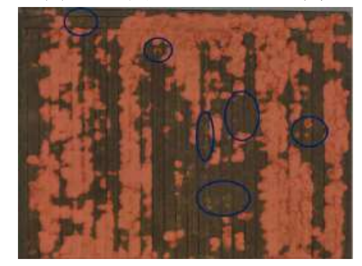

(f)迁移学习模型分割80 m图像

图4　小麦田倒伏面积提取结果举例

Fig. 4 Example of wheatland lodging area detection

3.2 不同研究区小麦倒伏面积提取平均精度对比

结合表3和图4，对比不同研究区的模型倒伏面积提取精度评价指标平均值，白湖的倒伏面积提取精度略优于舒城，舒城的倒伏面积提取精度优于郭河。其中，白湖倒伏面积提取结果的交并比、正确率、精确率、召回率和F_1-Score的平均数分别为89.25%、96.77%、93.89%、94.65%和94.18%，舒城倒伏面积提取结果的上述5个指标平均数分别为84.31%、96.68%、89.13%、93.34%和91.01%，郭河倒伏面积提取结果的5个指标平均分别为80.38%、89.87%、88.81%、89.77%和89.02%。

3.3 不同模型对80 m高度采集的小麦倒伏面积提取平均精度对比

评估对照模型、混合训练模型和迁移学习模型的性能是研究的核心工作，模型对80 m高度采集的图像倒伏面积提取的精度可以验证本研究提出的利用40 m图像预训练语义分割模型，再用少量80 m图像做迁移学习的策略的有效性。结合表3和图4，实验结果表明，对照模型对80 m高度采集的图像倒伏面积提取结果的交并比、正确率、精确率、召回率和F_1-Score的平均数分别为82.18%、93.30%、90.22%、90.13%和89.82%，混合训练模型的上述5个指标平均数分别为85.27%、94.83%、91.19%、92.68%和91.79%，迁移学习模型的上述5个指标平均数分别为85.37%、94.98%、91.30%、92.52%和91.84%。迁移学习模型的性能和混合训练模型的倒伏面积提取结果均高于对照模型，迁移学习模型的性能优于混合训练模型，其中，迁移学习模型的交并比、正确率、精确率和F_1-Score分别高于混合训练模型0.1%、0.15%、0.1%和0.05%，迁移学习模型召回率低于混合训练模型的0.16%。

3.4 模型算法效率分析

结合表4，经计算，对照模型、混合训练模型和迁移学习模型对80 m高度无人机图像倒伏面积提取的平均加权帧率分别为737.16、752.10和738.35 fps/m^2，而40 m图像对应的平均加权帧率分别为178.09、171.18和183.12 fps/m^2。可见使用80 m高度采集的图像进行倒伏监测的效率远高于40 m图像。

4　讨　论

本研究运用对比实验的方法，验证40 m图像预训练语义分割模型，再使用80 m图像迁移学习的策略，对倒伏面积提取模型精度提升的有效性。实验结果与预期一致，迁移学习模型精度优于对照

模型和混合训练模型。

在以往的研究中，学者们从遥感、农学、计算机、深度学习等角度开展了诸多倒伏面积提取算法研究，算法从理论层面实现了较高精度的倒伏面积提取。如Li等[24]基于特征工程算法提出的一种两步识别甘蔗倒伏区域的算法，实现了94%的倒伏面积提取精度，但是，研究区仅为2018年台风过后广西壮族自治区的一块甘蔗种植区。Yang等[25]基于深度学习自适应算法提出了一种水稻倒伏监测方法，运用浅层神经网络EDANet识别较高飞行高度采集图像中的倒伏区域，对疑似倒伏的区域降低无人机高度采集空间分辨率更高的图像，从而提升倒伏面积提取的精度和效率，实验结果虽然实现了99.25%的精度和36%倒伏面积提取效率的提升，但是图像训练集仅为台湾省台中市五峰区墨子盾公园1幅2017年采集的200 m飞行高度图像，测试集只有2019年5月21日采集的50 m飞行高度图像和5月23日200 m飞行高度图像。上述研究并没有遇到本研究解决的80 m图像数据量不足以支撑模型训练的问题，主要原因在于上述研究处理的数据仅仅为1或2年数据，使用特征工程和浅层神经网络即可实现较高的倒伏面积提取精度，但是，相关研究模型缺乏多年份数据验证，推广应用价值不足。本研究采集了4个年份3个研究区的大田小麦倒伏图像，受到田间非结构化环境影响，图像倒伏特征差异性较大，故需要采用参数量较多的Swin-Transformer模型进行训练[26,27]。Swin-Transformer模型训练将面临高飞行高度采集的图像数据量不足的难题，故提出了利用40 m图像预训练语义分割模型，再用少量80 m图像做迁移学习的策略解决上述问题。

精度和算法效率是衡量倒伏面积提取算法性能的2类重要评价标准，先前研究者采用各种图像处理和遥感领域的精度指标评价自身的研究成果。Liu等[28]基于无人机可见光和热成像图像，分别建立粒子群优化和支持向量机算法的水稻倒伏模型，实验结果表明，该方法对于倒伏和非倒伏区域的错误识别率均小于10%，估算倒伏面积R^2大于0.9。Song等[29]基于无人机遥感图像，提出一种图像融合和深度语义分割的向日葵倒伏识别方法，达到89.8%的精度。上述研究对于倒伏面积提取精度的评价指标较为笼统，仅仅使用总体识别精度作为指标，未能准确体现算法对于倒伏和非倒伏区域识别精度之间的关系。本研究引入交并比、正确率、精确率、召回率和F_1-Score指标和算法时间效率指标，系统地分析了算法的性能及运行效率。

由实验结果可知，迁移学习模型性能最佳，具体到3个研究区，白湖的倒伏面积提取精度优于舒城，舒城的倒伏面积提取结果优于郭河。造成上述差异的主要原因在于郭河研究区隶属安徽农业大学庐江科技园，承担了小麦品种选育、新农艺技术开发、农业病害防治技术研究等实验任务，麦田采取了条播的方式，种植了诸多品种小麦，种植密度大，倒伏小麦面积的比例较大，成熟晚期达到了64.99%，田间出现了较多半倒伏的小麦，造成了部分非倒伏小麦被误判为倒伏（图5），严重降低了召回率（郭河平均召回率为89.77%，低于白湖4.88%，低于舒城3.57%）和交并比（郭河平均交并比80.38%，低于白湖8.80%，低于舒城3.94%），郭河倒伏识别结果的正确率、精确度和F_1-Score指标也低于白湖和舒城。

5 结 论

为解决较高飞行高度图像数量无法满足语义分割模型训练要求，本研究提出迁移学习策略：在Swin-Transformer深度学习语义分割框架基础上，使用40 m高度采集的无人机图像预训练模型，再使用80 m高度采集的图像开展迁移学习训练；采用对比实验方法，比较对照模型、混合训练模型和迁移学习模型分割80 m高度采集图像的精度，评估模型性能。实验结果表明，迁移学习模型倒伏面积提取精度最高，交并比、正确率、精确率、召回率和F_1-Score指标平均数分别为85.37%、94.98%、91.30%、92.52%和91.84%，高于对照组模型1.08%～3.19%，平均加权帧率达到738.35 fps/m^2，高于40 m图像183.12 fps/m^2。本研究提出的策略是一种解决较高尺度图像数量无法满足语义分割模型训练要求问题的有效方法。相比先前研究成果[30,31]，本研究提出的方法可以在提升倒伏监测精度的同时，较大幅度地提升图像分割速度。在实际应用中，可以首先使用从多个研究区采集的包含多


(a)舒城半倒伏麦田


(b)白湖半倒伏麦田


(c)郭河半倒伏麦田

图5　舒城、白湖和郭河研究区半倒伏小麦

Fig. 5 Semi lodging wheat in the study areas of Shucheng, Baihu and Guohe

个小麦品种的大量40 m高度无人机图像进行预训练，再针对特定的研究区，采用少量小麦80 m高度图像进行迁移学习训练，获得针对性的倒伏检测模型。

受到实验条件的限制，研究区集中在安徽省中部，收集了无人机在40和80 m采集的麦田灌浆期和成熟期可见光图像。后续的研究将尝试使用更高高度（大于100 m）的无人机图像进行迁移学习训练，进一步提升倒伏面积提取效率。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 中华人民共和国国家统计局. 国家数据[EB/OL]. (2021-12-06)[2023-01-29]. http://www.stats.gov.cn/tjsj/.

[2] FLORES Paulo, 张昭. 基于无人机图像以及不同机器学习和深度学习模型的小麦倒伏率检测[J]. 智慧农业(中英文), 2021, 3(2): 23-34.
FLORES P, ZHANG Z. Wheat lodging ratio detection based on UAS imagery coupled with different machine learning and deep learning algorithms[J]. Smart agriculture, 2021, 3(2): 23-34.

[3] GUAN H X, HUANG J X, LI L, et al. A novel approach to estimate maize lodging area with PolSAR data[J]. IEEE transactions on geoscience and remote sensing, 2022, 60: 1-17.

[4] SINGH D, WANG X, KUMAR U, et al. High-throughput phenotyping enabled genetic dissection of crop lodging in wheat[J]. Frontiers in plant science, 2019, 10: ID 394.

[5] 韩东, 杨浩, 杨贵军, 等. 基于Sentinel-1雷达影像的玉米倒伏监测模型[J]. 农业工程学报, 2018, 34(3): 166-172.
HAN D, YANG H, YANG G J, et al. Monitoring model of maize lodging based on Sentinel-1 radar image[J]. Transactions of the Chinese society of agricultural engineering, 2018, 34(3): 166-172.

[6] WANG J J, GE H, DAI Q G, et al. Unsupervised discrimination between lodged and non-lodged winter wheat: A case study using a low-cost unmanned aerial vehicle[J]. International journal of remote sensing, 2018, 39(8): 2079-2088.

[7] HUANG X D, XUAN F, DONG Y, et al. Identifying corn lodging in the mature period using Chinese GF-1 PMS images[J]. Remote sensing, 2023, 15(4): ID 894.

[8] CHAUHAN S, DARVISHZADEH R, BOSCHETTI M, et al. Remote sensing-based crop lodging assessment: Current status and perspectives[J]. ISPRS journal of photogrammetry and remote sensing, 2019, 151: 124-140.

[9] 李宗南, 陈仲新, 任国业, 等. 基于Worldview-2影像的玉米倒伏面积估算[J]. 农业工程学报, 2016, 32(2): 1-5.
LI Z N, CHEN Z X, REN G Y, et al. Estimation of maize lodging area based on Worldview-2 image[J]. Transactions of the Chinese society of agricultural engineering, 2016, 32(2): 1-5.

[10] YANG H, CHEN E X, LI Z Y, et al. Wheat lodging monitoring using polarimetric index from RADARSAT-2 data[J]. International journal of applied earth observation and geoinformation, 2015, 34: 157-166.

[11] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]// 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, New Jersey, USA: IEEE, 2022: 9992-10002.

[12] HE X, ZHOU Y, ZHAO J Q, et al. Swin transformer embedding UNet for remote sensing image semantic segmentation[J]. IEEE transactions on geoscience and remote sensing, 2022, 60: 1-15.

[13] XU Y F, ZHOU S B, HUANG Y H. Transformer-based model with dynamic attention pyramid head for semantic segmentation of VHR remote sensing imagery[J]. Entropy, 2022, 24(11): ID 1619.

[14] KOH J C O, SPANGENBERG G, KANT S. Automated machine learning for high-throughput image-based plant phenotyping[J]. Remote sensing, 2021, 13(5): ID 858.

[15] ZHAO X, YUAN Y T, SONG M D, et al. Use of un-

manned aerial vehicle imagery and deep learning UNet to extract rice lodging[J]. Sensors, 2019, 19(18): ID 3859.
[16] YANG M D, TSENG H H, HSU Y C, et al. Semantic segmentation using deep learning with vegetation indices for rice lodging identification in multi-date UAV visible images[J]. Remote sensing, 2020, 12(4): ID 633.
[17] ZHAO J L, LI Z, LEI Y, et al. Application of UAV RGB images and improved PSPNet network to the identification of wheat lodging areas[J]. Agronomy, 2023, 13(5): ID 1309.
[18] ZHANG D Y, DING Y, CHEN P F, et al. Automatic extraction of wheat lodging area based on transfer learning method and deeplabv3+ network[J]. Computers and electronics in agriculture, 2020, 179: ID 105845.
[19] YU J, CHENG T, CAI N, et al. Wheat lodging extraction using Improved_Unet network[J]. Frontiers in plant science, 2022, 13: ID 1009835.
[20] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin, German: Springer, 2015: 234-241.
[21] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. arXiv: 1706.05587, 2017.
[22] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[EB/OL]. arXiv: 2103.14030, 2021.
[23] ZHANG G, YAN H F, ZHANG D Y, et al. Enhancing model performance in detecting lodging areas in wheat fields using UAV RGB Imagery: Considering spatial and temporal variations[J]. Computers and electronics in agriculture, 2023, 214: ID 108297.
[24] LI X H, LI X Z, LIU W, et al. A UAV-based framework for crop lodging assessment[J]. European journal of agronomy, 2021, 123: ID 126201.
[25] YANG M D, TSENG H H, HSU Y C, et al. Semantic segmentation using deep learning with vegetation indices for rice lodging identification in multi-date UAV visible images[J]. Remote sensing, 2020, 12(4): ID 633.
[26] BISWAL S, CHATTERJEE C, MAILAPALLI D R. Damage assessment due to wheat lodging using UAV-based multispectral and thermal imageries[J]. Journal of the Indian society of remote sensing, 2023, 51(5): 935-948.
[27] GAO L, LIU H, YANG M H, et al. STransFuse: Fusing swin transformer and convolutional neural network for remote sensing image semantic segmentation[J]. IEEE journal of selected topics in applied earth observations and remote sensing, 2021, 14: 10990-11003.
[28] LIU T, LI R, ZHONG X C, et al. Estimates of rice lodging using indices derived from UAV visible and thermal infrared images[J]. Agricultural and forest meteorology, 2018, 252: 144-154.
[29] SONG Z S, ZHANG Z T, YANG S Q, et al. Identifying sunflower lodging based on image fusion and deep semantic segmentation with UAV remote sensing imaging[J]. Computers and electronics in agriculture, 2020, 179: ID 105812.
[30] LI G A, HAN W T, HUANG S J, et al. Extraction of sunflower lodging information based on UAV multi-spectral remote sensing and deep learning[J]. Remote sensing, 2021, 13(14): ID 2721.
[31] TIAN M L, BAN S T, YUAN T, et al. Assessing rice lodging using UAV visible and multispectral image[J]. International journal of remote sensing, 2021, 42(23): 8840-8857.

Identification Method of Wheat Field Lodging Area Based on Deep Learning Semantic Segmentation and Transfer Learning

ZHANG Gan[1], YAN Haifeng[1], HU Gensheng[1*], ZHANG Dongyan[1,2], CHENG Tao[1,2], PAN Zhenggao[1,3], XU Haifeng[1,3], SHEN Shuhao[1,3], ZHU Keyu[1]

(1. *National Engineering Research Center for Argo-Ecological Big Data Analysis &Application*, *Anhui University*, *Hefei* 230039, *China*; 2. *College of Mechanical and Electronic Engineering*, *Northwest A&F University*, *Yangling* 712100, *China*; 3. *School of Information Engineering*, *Suzhou University*, *Suzhou* 234000, *China*)

Abstract:

[Objective] Lodging constitutes a severe crop-related catastrophe, resulting in a reduction in photosynthesis intensity, diminished nutrient absorption efficiency, diminished crop yield, and compromised crop quality. The utilization of unmanned aerial vehicles (UAV) to acquire agricultural remote sensing imagery, despite providing high-resolution details and clear indications of crop lodging, encoun-

ters limitations related to the size of the study area and the duration of the specific growth stages of the plants. This limitation hinders the acquisition of an adequate quantity of low-altitude remote sensing images of wheat fields, thereby detrimentally affecting the performance of the monitoring model. The aim of this study is to explore a method for precise segmentation of lodging areas in limited crop growth periods and research areas.

[Methods] Compared to the images captured at lower flight altitudes, the images taken by UAVs at higher altitudes cover a larger area. Consequently, for the same area, the number of images taken by UAVs at higher altitudes is fewer than those taken at lower altitudes. However, the training of deep learning models requires huge amount supply of images. To make up the issue of insufficient quantity of high-altitude UAV-acquired images for the training of the lodging area monitoring model, a transfer learning strategy was proposed. In order to verify the effectiveness of the transfer learning strategy, based on the Swin-Transformer framework, the control model, hybrid training model and transfer learning training model were obtained by training UAV images in 4 years (2019, 2020, 2021, 2023)and 3 study areas(Shucheng, Guohe, Baihe) under 2 flight altitudes (40 and 80 m). To test the model's performance, a comparative experimental approach was adopted to assess the accuracy of the three models for segmenting 80 m altitude images. The assessment relied on five metrics: intersection of union (IoU), accuracy, precision, recall, and F_1-score.

[Results and Discussions] The transfer learning model shows the highest accuracy in lodging area detection. Specifically, the mean IoU, accuracy, precision, recall, and F_1-score achieved 85.37%, 94.98%, 91.30%, 92.52% and 91.84%, respectively. Notably, the accuracy of lodging area detection for images acquired at a 40 m altitude surpassed that of images captured at an 80 m altitude when employing a training dataset composed solely of images obtained at the 40 m altitude. However, when adopting mixed training and transfer learning strategies and augmenting the training dataset with images acquired at an 80 m altitude, the accuracy of lodging area detection for 80 m altitude images improved, inspite of the expense of reduced accuracy for 40 m altitude images. The performance of the mixed training model and the transfer learning model in lodging area detection for both 40 and 80 m altitude images exhibited close correspondence. In a cross-study area comparison of the mean values of model evaluation indices, lodging area detection accuracy was slightly higher for images obtained in Baihu area compared to Shucheng area, while accuracy for images acquired in Shucheng surpassed that of Guohe. These variations could be attributed to the diverse wheat varieties cultivated in Guohe area through drill seeding. The high planting density of wheat in Guohe resulted in substantial lodging areas, accounting for 64.99% during the late mature period. The prevalence of semi-lodging wheat further exacerbated the issue, potentially leading to misidentification of non-lodging areas. Consequently, this led to a reduction in the recall rate (mean recall for Guohe images was 89.77%, which was 4.88% and 3.57% lower than that for Baihu and Shucheng, respectively) and IoU (mean IoU for Guohe images was 80.38%, which was 8.80% and 3.94% lower than that for Baihu and Shucheng, respectively). Additionally, the accuracy, precision, and F_1-score for Guohe were also lower compared to Baihu and Shucheng.

[Conclusions] This study inspected the efficacy of a strategy aimed at reducing the challenges associated with the insufficient number of high-altitude images for semantic segmentation model training. By pre-training the semantic segmentation model with low-altitude images and subsequently employing high-altitude images for transfer learning, improvements of 1.08% to 3.19% were achieved in mean IoU, accuracy, precision, recall, and F_1-score, alongside a notable mean weighted frame rate enhancement of 555.23 fps/m^2. The approach proposed in this study holds promise for improving lodging monitoring accuracy and the speed of image segmentation. In practical applications, it is feasible to leverage a substantial quantity of 40 m altitude UAV images collected from diverse study areas including various wheat varieties for pre-training purposes. Subsequently, a limited set of 80 m altitude images acquired in specific study areas can be employed for transfer learning, facilitating the development of a targeted lodging detection model. Future research will explore the utilization of UAV images captured at even higher flight altitudes for further enhancing lodging area detection efficiency.

Key words: lodging detection; agricultural remote sensing; UAV image; transfer learning; semantic segmentation; Swin-Transformer

Foundation items: University Research Project of Anhui Provincial Department of Education (Natural Science)(2023AH052246); Suzhou University Doctoral Research Foundation (2021BSK043); National Natural Science Foundation of China (42271364)